PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68

A2

(11) International Publication Number: WO 00/26415

(43) International Publication Date: 11 May 2000 (11.05.00)

(21) International Application Number: PCT/US99/25792

(22) International Filing Date: 2 November 1999 (02.11.99)

(30) Priority Data:
60/106,701 2 November 1998 (02.11.98) US

(71) Applicant: THE BRIGHAM & WOMEN'S HOSPITAL, INC. [US/US]; 75 Francis Street, Boston, MA 02115 (US).

(72) Inventors: FLETCHER, Jonathan; 122 Pleasant Street #3, Brookline, MA 02246 (US). XIAO, Sheng; 400 Brookline Avenue #10F, Boston, MA 02215 (US).

(74) Agent: LOCKHART, Helen, C.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

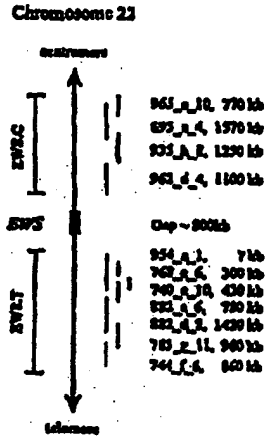
(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

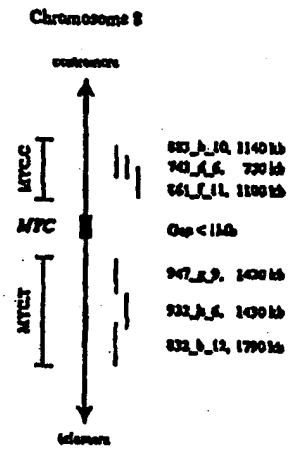
Published
Without international search report and to be republished upon receipt of that report.

(54) Title: A METHOD FOR GENOMIC SUBSTRACTIVE HYBRIDIZATION

(57) Abstract

The invention is a method for genomic subtractive hybridization. Specific nucleic acid sequences are removed from a sample of nucleic acid sequences by specifically hybridizing the sequences to a complementary nucleic acid sequence bound to a target molecule such as biotin. The target molecule is then contacted with a binding partner such as avidin and separated from the sample of nucleic acid sequences. As the target is separated from the sample the hybridized nucleic acid sequences are also removed from the sample. The method preferably involves the removal of repetitive nucleic acid sequences from a nucleic acid sample to generate a library of probes that are substantially free of repetitive nucleic acid sequences.





FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

A METHOD FOR GENOMIC SUBTRACTIVE HYBRIDIZATION

Field Of The Invention

The present invention relates to a method for genomic subtractive hybridization. The invention also relates to methods for generating probes for in situ hybridization techniques. In particular, the methods for genomic hybridization are useful for generating probes for techniques such as in situ hybridization.

Background Of The Invention

Chromosomes within a living cell encompass all of the DNA of a particular organism. The number and structure of chromosomes within a cell are indicative of various normal and abnormal developmental traits. For instance, the presence of a particular chromosome type or number of a particular chromosome may be indicative of an abnormal condition. Humans having three copies of chromosome 18 (trisomy) develop a fatal disorder causing death within one year of birth and subjects having three copies of chromosome 21 (trisomy 21) develop Downs syndrome. In general chromosome abnormalities can result from extra or missing individual chromosomes, extra or missing portions of a chromosome, or chromosomal rearrangements, including translocations, deletions, and inversions. The trisomies discussed above involve addition of chromosomal material. A translocation involves the transfer of a piece of a chromosome to another chromosome. An example of a disorder involving an exchange of chromosomal material is chronic myelogenous leukemia which involves a translocation of chromosomal material from chromosome 9 to chromosome 22. An inversion involves a reversal in polarity of a chromosomal segment. Dicentrics produce a chromosome with two centromeres.

Characteristic chromosome abberations have been described in a wide range of tumors. Specific oncogene and tumor suppressor gene targets affected by these chromosomal abnormalties have been characterized in some tumors but most of them remain to be studied. A major goal in studying human cancer chromosome and gene aberrations is to elucidate biological pathways responsible for neoplastic transformation. Several such pathways have already been identified through characterization of particular cancer chromosome aberrations (1, 3). Another goal in cancer chromosome evaluation is the identification and validation of novel diagnostic and prognostic markers (4, 5). Cytogenetic markers have utility as histological adjuncts, and methods that delineate diagnostic chromosomal aberrations or

associated molecular changes are becoming increasingly important in experimental pathology (6-8).

Chromosomal abnormalities are generally detected in biological samples using standard methods for analyzing karyotypes. A karyotype defines the number and morphology of chromosomes in an individual or a related group of individuals. The number can be determined as the total chromosome number or the copy number of individual chromosome types. Chromosomal morphology can be determined for example, by measuring length of the chromosome or the centromeric index of the chromosome. Typically, techniques based on chemical staining have been used to produce bands on the chromosome that allow for the identification of each chromosome type. The banding technique has been used for years to identify chromosomal abnormalities such as translocations and inversions. (Latt, Optical Studies of Metaphase Chromosome Organization, Annual Review of Biophysics and Bioengineering, v. 5, p.1-37 (1976)).

The analysis of human tumors by the banding technique, however, has been difficult because primary human tumor cells generally are difficult to culture and there are often too few cells in metaphase to accurately band metaphase chromosomes. In order to accomplish the banding, the cells generally need to be stimulated to undergo cellular proliferation using mitogens. Tumor cells in a solid tumor often cannot even be stimulated to undergo cellular proliferation.

In situ hybridization (ISH) is an analytic method that enables evaluation of numerical and structural chromosome aberrations in both fresh and archival tumor and other tissue specimens (9-13) and has extended the capabilities of conventional tumor and other tissue karyotyping (14-17). ISH analysis using chromosome specific nucleic acid probes solve the problems associated with karyotyping of tumor cells with the banding technique by allowing the analysis of interphase nuclei, and thus avoiding the need to prepare metaphase chromosomes. Several advantages of ISH include applicability to interphase cell populations and the ability to demonstrate unambiguously chromosome rearrangements that target specific loci of interest (19-20). Several technical innovations have permitted the application of ISH in genome-wide screening studies, including comparative genomic hybridization (CGH) (21), combinatorial multifluor ISH (22), and multicolor spectral karyotyping (23). Each of these methods provides a broad perspective complementary to that provided by chromosome banding methods.

ISH has been used to identify the location of individual genes or other well defined nucleic acid sequences on chromosomes. Single copies of unique sequence probes have been used to identify the location of a particular gene on a chromosome. High background levels, however, often prevent reliable detection of small target sites.

Chromosome specific libraries have also been developed to generate probes useful for ISH to detect chromosome specific abnormalities. One method for generating chromosome specific libraries involves the use of flow cytometry to isolate a pure preparation of a chromosome from a pool of many individual chromosomes. The isolated chromosome may then be amplified by polymerase chain reaction (PCR) to manufacture a library (Chang et al., Genomics, v. 12, p. 307-312, 1992 and Boschman et al., Genes, Chromosomes, and Cancer, v. 6, p. 10-16, 1993). Additionally, chromosome specific libraries have been prepared from somatic cell hybrids (e.g., a rat-human hybrid cell line) which contains a few distinct human chromosomes. The chromosomes can be isolated and generated into a library without the requirement of a flow cytometry step. More recently, these types of libraries have been produced from subchromosomal regions using yeast, artificial chromosomes (YACs) that correspond to a region of interest on a particular chromosome (Lengauer et al., Hum. Mol. Genet., v. 2, p. 505-512, 1993).

Summary of the Invention

The present invention relates to methods and products for genomic subtractive hybridization. The methods are useful in particular for developing unique-sequence genomic probes suitable for ISH procedures. As described above, ISH is commonly used for pathological screening purposes to detect chromosomal defects. The genomic subtractive hybridization methods of the invention produce stable probes for ISH depleted of non-specific repeat sequences and thus produce less background than traditional methods of ISH. The probes produced according to these methods also are a stable resource of ISH DNA that can be labeled by many different types of methods including PCR, random priming or nick translation.

In one aspect the invention is a method for genomic subtractive hybridization. The method involves the steps of hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in a nucleic acid sample, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule, and selectively removing the chemically modified oligonucleotide probe and complementary

nucleic acid by selectively contacting the target molecule with a binding partner to produce binding partner/target conjugates and separating the binding partner and binding partner/target conjugates from the nucleic acid sample.

In one embodiment the target molecule is selected from the group consisting of avidin, biotin, fluorescein isothiocyanate (FITC), anti-FITC, antigen, and antibodies. Preferably the target molecule is biotin or FITC. In one embodiment the binding partner is immobilized on a support. In another embodiment the support is a bead and wherein the binding partner/target conjugates are separated from the nucleic acid sample by chromatography.

The oligonucleotide may have any specific nucleic acid sequence but preferably the oligonucleotide is a repetitive nucleic acid. In one embodiment the repetitive nucleic acid is isolated from YAC DNA or BAC DNA. In a preferred embodiment the chemically modified oligonucleotide probe is prepared by amplifying the repetitive nucleic acid using PCR and a primer attached to the target molecule.

The nucleic acid sample is a genomic nucleic acid sample in another embodiment. Preferably the genomic nucleic acid sample is obtained from a YAC, BAC, PAC, or P1 clone.

According to another aspect of the invention a library of nucleic acid probes for use in in situ hybridization methods is provided. The library includes a heterogenous mixture of labeled nucleic acid probes that are substantially complementary to unique nucleic acid fragments and are substantially free of repetitive nucleic acid sequences, and which are produced by the process of: (a) obtaining genomic nucleic acid fragments; (b) amplifying the genomic nucleic acid fragments; (c) hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in the genomic nucleic acid fragments, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule; (d) selectively removing the chemically modified oligonucleotide probe and complementary nucleic acid by selectively contacting the target molecule with a binding partner and separating the binding partner and binding partner/target conjugates from the genomic nucleic acid fragments; and (e) labeling the genomic nucleic acid fragments with a label. Preferably the genomic nucleic acid fragments are labeled with biotin, digoxigenin, or a fluorescent molecule.

In one embodiment the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of the unique nucleic acid fragments in a genomic DNA population. In another embodiment the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of the unique nucleic acid fragments in a chromosome of a genomic DNA population. In yet another embodiment the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of the unique nucleic acid fragments in a subregion of a chromosome of a genomic DNA population.

In another aspect the invention is a method for performing in situ hybridization. The method includes the steps of hybridizing the library of labeled probes of the invention to a biological sample fixed on a surface, removing un-hybridized probe, and detecting a signal from the hybridized probe to identify nucleic acid sequences present in the biological sample.

According to another aspect of the invention a mixture of probes for genomic subtractive hybridization is provided. The probes of the mixture are isolated chemically modified oligonucleotide probes, wherein the chemically modified oligonucleotide probes are oligonucleotides associated with a target molecule selected from the group consisting of biotin, FITC, and wherein the oligonucleotide probes are a mixture of oligonucleotides complementary to repetitive nucleic acid sequences and yeast nucleic acid sequences. Preferably the target molecule is biotin.

In one embodiment the repetitive nucleic acid is isolated from YAC, BAC, PAC, or P1 clone. In another embodiment the chemically modified oligonucleotide probe is prepared by amplifying the repetitive nucleic acid using PCR and a primer attached to the target molecule.

Each of the limitations of the invention can encompass various embodiments of the invention. It is, therefore, anticipated that each of the limitations of the invention involving any one element or combinations of elements can be included in each aspect of the invention.

Brief Description Of The Drawings

Figure 1 is a chromosomal map depicting regions of chromosome 8 and chromosome 22.

Figure 2 is a dot blot of subtracted and unsubtracted nucleic acid probes probed with Cot-1 (panel A) or total yeast genomic DNA (Panel B).

Brief Description Of The Sequences

SEQ. ID. NO. 1 is a primer having the sequence 5'CTGAGCGGAATTCGTGAGACC (T-1)

SEQ. ID. NO. 2 is a primer having the sequence 5'GGTCTCACGAATTCCGC TCAGTT (T-2)

SEQ. ID. NO. 3 is a primer having the sequence 5'AATTCTTGCGCCTTAAACC AAC (D-40)

SEQ. ID. NO. 4 is a primer having the sequence 5'GTTGGTTTAAGGCGC AAG (D-41).

SEQ. ID. NO. 5 is a primer having the sequence 5'AATTCTTGCGCCTTAA ACCAAC (D-40B).

Detailed Description Of The Invention

The invention relates to a novel approach for expanding the role of *in situ* hybridization in diagnostic and research applications. Repetitive DNA sequences have been used in ISH to characterize various human chromosomes (Devillee et al., Cytogenet. Cell Genet., v. 41, p. 193-201 1986), for instance, to detect the presence of multiple loci on a chromosome. ISH probes derived from pericentromeric α-satellite repeat sequences have been used extensively to evaluate numerical chromosome aberrations. Many molecular cytogenetic research and diagnostic questions, however, are best addressed using unique sequence probes to particular chromosome regions. Large-insert clones, e.g., BAC, PAC, P1, and YAC clones, are suitable for molecular cytogenetic applications, including those using paraffin-embedded material. The fluorescence signals obtained using large-insert clones are often as bright as those obtained with α-satellite probes, particularly when the human genomic inserts are rich in unique sequences. The invention provides new methods and related products for preparing libraries of probes that are rich in unique sequences.

When used with the genomic subtractive hybridization methods of the invention these large probes are also suitable for bright-field studies, using either peroxidase/alkaline phosphatase, or sequential peroxidase detection strategies. In addition to the potential for excellent ISH signal intensity, a major advantage of mega-YACs is the wealth of physical mapping data accessible via the Internet at websites such as those of the Whitehead Institute and CEPH. These data facilitate selection of mega-YAC clones in a region of interest. As described in the Examples below multiple mega-YACs in the EWS and MYC regions have

been tested and clones were selected with superior ISH signals for inclusion in the final contigs (Figure 1). The probes (YAC contigs spanning 1.5 to 5.0 megabases) were treated by the subtractive hybridization methods of the invention to produce unique-sequence rich DNA.

The genomic subtractive hybridization approach of the invention accomplishes at least two objectives. The first is the removal of repetitive-sequences, and the second is the creation of a library of DNA fragments that can be manipulated readily for large-scale DNA preparation and labeling. Repetitive sequences in prior art ISH probes are generally competed through pre-annealing with total genomic or repeat-sequence enriched DNA.[11,56] Pre-annealing can be an effective approach, but this step is both expensive and time consuming. In addition, some labeled repetitive sequences will likely come in contact with the slide. It is fundamentally desirable to remove these sequences from the probe altogether, thus eliminating a potential source of nonspecific fluorescence. The methods of the invention avoid these problems and thus accomplish the above-stated objectives.

In one aspect the invention is a method for genomic subtractive hybridization. The method involves the steps of hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in a nucleic acid sample, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule, and selectively removing the chemically modified oligonucleotide probe and complementary nucleic acid by selectively contacting the target molecule with a binding partner to produce binding partner/target conjugates and separating the binding partner and binding partner/target conjugates from the nucleic acid sample.

A "chemically modified oligonucleotide probe" is an oligonucleotide associated with a target molecule. The oligonucleotide can have any nucleic acid sequence that is desirable. The purpose of the oligonucleotide is to hybridize to and remove a complementary nucleic acid from a mixture of nucleic acid sequences. Therefore it is desirable to have an oligonucleotide which has a nucleic acid sequence that is complementary to a nucleic acid sequence that will be removed. For instance when hybrid cells are used to generate chromosomes it is desirable to remove all non-human chromosomes. In this case the oligonucleotide would have a sequence complementary to the non-human sequences, such as yeast DNA. Preferably the oligonucleotide is a repetitive nucleic acid sequence (or a mixture of repetitive nucleic aid sequences and yeast DNA) which can be used to remove human repetitive sequences (repetitive sequences and yeast) from a library of oligonucleotide probes

to reduce background levels in an ISH procedure. The oligonucleotide may be modified or unmodified. For instance the oligonucleotide may have a partial or complete phosphorothioate backbone.

As used herein "repetitive nucleic acid sequences" are nucleic acid sequences within a genome which encompass a series of nucleotides which are repeated many times, often in tandem arrays. The repetitive sequences can occur in the genome in multiple copies ranging from two to hundreds of thousands of copies and may be clustered or interspersed on one or more chromosomes throughout a genome. Repetitive nucleic acid sequence fragments when present in a pool of nucleic acid fragments which are denatured and then rehybridized, will rehybridize at a more rapid rate than the unique sequences. Unique sequences are those sequences having at least 12 nucleotides and which are present only once within an entire genome. Although the repetitive nucleic acid sequences are present throughout the genome a large number of the repetitive nucleic acid sequences are located at the centromere of each chromosome.

The human genome includes several families of repetitive nucleic acid sequences. For instance, alpha satellite nucleic acid sequences have variable numbers of tandem sequences of approximately 170 base pairs long. Satellite 2 nucleic sequences have repeat units of approximately 26 base pairs long. The repetitive nucleic acid sequences within a family are highly homologous. These families also include subfamilies of repetitive sequences which have been produced by substitutions, deletions, and/or insertions or repetitive nucleic acid sequences. The degree of variation depends on the particular subfamily. As used herein the repetitive nucleic acid sequence refers to families and subfamilies which are capable of hybridizing with the dominate form of the family under normal stringency conditions of temperature, concentration, and time.

The oligonucleotide of the chemically modified oligonucleotide probe is associated with a target molecule. As used herein the term "associated" refers to a covalent interaction between the target molecule and the oligonucleotide.

The genomic nucleic acid fragments are hybridized with a chemically modified oligonucleotide probe which is complementary to some of the nucleic acid sequences in the genomic acid fragments. The hybridization step is preferably performed in a solution under standard conditions. For instance, the oligonucleotide probe and the genomic nucleic acid

fragments should be denatured at elevated temperatures and then incubated at approximately 65°C for 24-48 hours in the presence of carrier molecules such as yeast tRNA.

A "target molecule" as used herein is a molecule which can be covalently attached to an oligonucleotide and which has a known binding partner. Preferably the target molecule is selected from a group consisting of biotin, avidin, FITC, anti-FITC, antigen, and antibodies. A binding partner is a molecule which interacts with a target molecule. The interaction is strong enough such that the binding partner can be used to remove target from a solution. Many binding partner/target pairs are well known in the art, e.g., biotin/avidin, FITC/anti-FITC.

The target molecule is caused to interact with a binding partner. The binding partner is removed from the mixture along with any binding partner/target conjugates which have formed. The remaining genomic nucleic acid fragments are then purified and labeled and can be used in ISH methods. The hybridization and selective removal steps can be repeated several times to assure removal of all complementary nucleic acid sequences. The labeling step may be performed by any method known in the art. The nucleic acid fragments may be labeled, for instance, using chemical modification by substituting derivitized bases, by forming adducts, which can be detected by immunochemical stains, or the addition of extrinsic labels such as radioactivity and fluorescence molecules.

The binding partner is preferably immobilized on a solid support. A solid support as used herein refers to any solid material to which a binding partner can be immobilized. Solid supports, for example, include but are not limited to membranes, e.g., nitrocellulose or nylon, a bead, e.g., a magnetic bead or particle, or polymers such as polystyrene. For instance, the binding partner may be immobilized on a bead and the binding partner/target conjugate can be separated by chromatography.

The invention in another aspect is a library of nucleic acid probes for use in ISH methods. The library of nucleic acid probes is a heterogeneous mixture of labeled nucleic acid probes that are substantially complementary to unique nucleic acid fragments and are substantially free of repetitive nucleic acid sequences, and which are produced by the process of obtaining genomic nucleic acid fragments, amplifying the genomic nucleic acid fragments, hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in the genomic nucleic acid fragments, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule, selectively

removing the chemically modified oligonucleotide probe in complementary nucleic acid by selectively contacting the target molecule with a binding partner to produce a binding partner/target conjugate and separating the binding partner and binding partner/target conjugates from the genomic nucleic acid fragments, and labeling the genomic nucleic acid fragments with a label. A library as used herein refers to a plurality of nucleic acid molecules.

A "heterogenous mixture of labeled nucleic acid probes" as used herein is a plurality of nucleic acid fragments which are complementary to unique nucleic acid sequences. A "unique nucleic acid sequence" as used herein is a nucleic acid sequence of at least 12 nucleotides in length and which is present only once within an entire genome. A complementary nucleic acid is one which is capable of base pairing with a nucleic acid under stringent conditions. Hybridization conditions are described, for instance, in Sambrook, et al., Molecular Cloning: A Laboratory Manual, second edition, V 1-3, Cold Springs Harbor Laboratory Press: Cold Spring Harbor, N.Y., 1989. The preferred length of the probe is at least 15 nucleotides. The number of nucleic acid fragments within the mixture may vary widely depending upon their use. For instance, if the mixture is used as a probe for ISH to identify the presence of certain specific nucleic acid sequences within a subregion of a chromosome, then the number of nucleic acid fragments will be less than if the ISH is performed to identify nucleic acids in a plurality of chromosomes. The number of fragments will also depend upon various experimental conditions such as the unit length of nucleic acid per unit volume that can be maintained in solution. Such parameters are well known to those of ordinary skill in the art.

The heterogenous mixture of labeled nucleic acid probes is substantially free of repetitive nucleic acid sequences. "Substantially free" as used herein refers to elimination of at least 95% of known repetitive sequences.

The step of obtaining genomic nucleic acid fragments may be performed by many methods known in the art. Preferably, the genomic nucleic acid fragments are obtained from isolated chromosome specific DNA. If the heterogeneous mixture of labeled nucleic acid probes are made from a single chromosome, then that chromosome is isolated and DNA is extracted from the isolated chromosomes. Individual chromosomes can be isolated by a variety of methods known in the art. One method involves the isolation of human chromosomes from hybrid cell lines formed between human cells and non-human cells. As

these hybrid cells are propagated many of the human chromosomes are lost and thus cells can be formed which contain a single human chromosome. Another method for isolating chromosomes is by direct flow sorting of metaphase cells. Such methods can be accomplished using commercially available equipment such as fluorescence-activated sorting instruments, e.g., Becton, Dickinson & FACS-II. DNA is then extracted from the isolated chromosomes by routine methods known in the art, e.g., Sambrook, et al., Molecular Cloning: A Laboratory Manual, second edition, V 1-3, Cold Springs Harbor Laboratory Press: Cold Spring Harbor, N.Y., 1989.

A preferred means for isolating individual chromosomes and regions thereof is through the use of yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), or P1-derived artificial chromosomes (PACs) containing specific human chromosomes. Many YAC and BAC clones are well known in the art and specific regions of chromosomes within these YAC, BAC, and PAC clones can be selected based on published maps as well as data obtained from websites such as the Whitehead Institute Center for genome research and CEPH-Genethon websites (http:\\www.genome.wi.mit.edu and http:\\www.cephb.fr, respectively). The BAC , YAC, and PAC DNA inserts can be isolated, and separated from their respective chromosomes, by methods that are well known in the art.

In preferred embodiments the YAC clone inserts can be isolated from yeast chromosomes by pulsed field gel electrophoresis, and specific size fractions may then be removed from the gels and purified. Pulsed field gel electrophoresis (***PFGE***) involves the resolution of large sized DNA molecules by periodically changing the electric field pattern during electrophoresis. The changes in field pattern reorient the DNA molecules and the separating medium, thus improving DNA separation. PFGE techniques have been extensively described in the prior art, such as US patent 5,135,628, which is incorporated by reference. The use of PFGE to isolate the YAC clone inserts from the yeast chromosomes provides unexpectedly good results. The DNA, in yeast cells containing YACs, is about 97% yeast chromosomal DNA (i.e. the normal yeast chromosomes) whereas 3% is the YAC. The YACs contain 1 Megabase-length human inserts (one million base pairs of DNA) whereas the yeast chromosomal DNA is about 30 megabases in total length. It was found according to the invention that when DNA is isolated from the yeast using PFGE the 1 Mb YAC is separated from similarly -sized individual yeast chromosomes. It was discovered that 500 ng of non-PFGE purified, subtracted, YAC probe is required to perform FISH analysis on one

specimen but that only 20 ng of PFGE-purified, subtracted, YAC probe is required to do FISH on one specimen. Thus the PFGE-purified material is 25-fold more potent because the irrelevant yeast chromosomal "filler" DNA sequences have been eliminated and only the actual human insert (probe) remains.

Once isolated, the human chromosomal DNA can be amplified. The term "amplification" as used herein refers to a method of producing complementary nucleic acid, such as PCR and other well-known methods. The step of amplifying the genomic nucleic acid fragments is preferably carried out using PCR. For instance, the nucleic acid fragments can be blunt ended and ligated to an adapter, such as the T-1/T-2 adapter, and then amplified using PCR. The amplified fragments can then be purified using techniques known in the art or commercially available kits such as QIA-quick PCR purification kits (QIAGEN).

Once the chromosomal DNA is prepared the subtractive hybridization methods of the invention can be used to synthesize ISH probes. Subtractive hybridization has been used in the prior art primarily for analysis of differences between two DNA populations.[47-51] Subtraction enables enrichment of the population of interest, (also referred to as tracer DNA), for sequences absent or substantially under represented in the nucleic acid population (also referred to as the driver). The methods according to the invention preferably utilize subtraction to remove repetitive sequences, while preserving unique sequences, in the mega-YAC contigs. The physical basis of the subtraction lies in the kinetics of DNA annealing in solution.[50] A molar excess of biotinylated repetitive DNA is used causing the denatured repetitive sequences in the nucleic acid sample population are far more likely to anneal to their counterparts in the biotinylated population rather than to each other. Biotinylated molecules, including binding partner/target hybrids, are subsequently removed from solution using a matrix of avidin-conjugated beads. The method maximizes both quantitative removal of repetitive sequences and retention of sequence complexity in the tracer.

We demonstrate in the Examples below that subtracted mega-YAC probes may be amplified through at least four rounds of PCR without diminishing the ISH signal intensity. Each 25-cycle round of PCR amplifies the starting material 500- to 1000-fold. Additional amplification is then achieved using random ocismer priming[45] or PCR to incorporate labeled nucleotides. Hence, subtraction converts 250 ng of adapter-ligated tracer DNA into a virtually permanent resource that can serve to generate kilogram amounts of repeat-sequence-depleted probes for ISH. This approach is particularly advantageous in the case of CEPH

mega-YACs, which are known to be unstable in the strain of yeast in which they were cloned.[57-58] CEPH mega-YACs often develop large deletions due to recombination and selection events. By converting contigs of mega-YAC clones into adapter-ligated libraries, we have created stable and readily available probes.

There are several well established techniques for production of large-insert ISH probes. These include *Alu*-PCR and DOP-PCR, which are rapid and universally applicable methods for amplification of YAC or BAC human inserts.[50-61] *Alu*-PCR is performed using *Alu*-sequence primers. An advantage in this approach is that *Alu*-containing, ie., human, sequences are amplified. However, the amplified sequences are limited primarily to those situated between closely neighboring, and appropriately oriented. *Alu* repeats, thus biasing the final pool of amplified DNA fragments, DOP-PCR is performed using degenerate oligonucleotide primers. This method, although nonselective for human versus bacterial or yeast sequences, may enable a more complex and representative amplification of the human insert than does *Alu*-PCR. Both *Alu*-PCR and DOP-PCR are effective and straight-forward methods requiring substantially less up-front effort than subtraction. However, neither approach is designed to generate probes depleted of repetitive sequences. In our experience, *Alu*-PCR and DOP-PCR YAC probes are generally associated with a lower signal-to-noise ratio, despite Cot-1 pre-annealing, than subtracted probes. DOP-PCR chromosome painting probes, on the other hand, have been extremely successful.[54-62]

The *EWS* and *MYC* probe sets described in the Examples are particularly effective in a screening mode. *EWS* translocations are found in Ewing's sarcoma, clear-cell sarcoma, desmoplastic small-round-cell tumor, extraskeletal myxoid chondrosarcoma, and (infrequently) myxoid liposarcoma. At least nine different partner genes participate in the translocation-related *EWS* fusion oncogenes in these tumors.[26-34] The *EWS* ISH screening approach allows efficient evaluation of an *EWS* rearrangement in any of the aforementioned tumors. In cases with abnormal ISH patterns, the specific fusion oncogene can then be established using appropriate oligonucleotide primers, by reverse transcription PCR. This screening approach also enables localization of previously uncharacterized translocation partners. *MYC* translocations are also particularly amenable to ISH detection. Translocation breakpoints, in HIV-associated and endemic Burkitt lymphomas, are often 100 to 300 kb upstream of *MYC*.[64-65] whereas translocation breakpoints in most sporadic Burkitt lymphomas involve *MYC* exon 1 or intron 1.[64,65] However, 10% to 20% of sporadic Burkitt lymphomas

contain arrangements of κ or λ light chain loci, and these cases typically have translocation breakpoints 200 to 300 kb downstream of *MYC*.[66] A *MYC* ISH probe set with a 500-kb gap on either side of the gene (Figure 1) was designed according to the invention such that virtually all *MYC* translocations, whether upstream, intragenic, or downstream, are detected.

"ISH" as used herein is a method for detecting and localizing nucleic acids within a cell or tissue preparation. The method provides both quantitative and spacial information concerning the nucleic acid sequences within an individual cell or chromosome. ISH has been commonly been used in the areas of prenatal genetic disorder diagnosis, molecular cytogenetics, to detect gene expression and overexpression, to identify sites of gene expression, to map genes, to localize target genes and to identify various viral and microbial infections, tumor diagnosis in vitro fertilization analysis, analysis of bone marrow transplantation and chromosome analysis. The technique involves the use of labeled nucleic acid probes which are hybridized to a chromosome or mRNA in cells that are immobilized on a slide. The probes can be labeled with fluorescent molecules, in a procedure known as fluorescent *In situ* hybridization (FISH), see, e.g., Kuo et al., Am. J. Hum. Genet., v. 49, p. 112-119, 1991). An example of ISH and FISH methods is provided in U.S. Patent No. 5, 750,340, issued to Kim et al.

ISH methods involve the fixation of tissue or biological samples on a surface, prehybridization treatment to increase the accessability of target DNA in the sample and to reduce non-specific binding, hybridization of the labeled library of nucleic acid probes to the DNA, post-hybridization washes to remove unbound probe, and detection of the hybridized probes. Each of these steps is well known in the art and has been performed under many different experimental conditions. Set forth below are some commonly used methods for performing ISH. The materials of the invention and the genomic subtractive hybridization steps of the invention can be used with any ISH procedure.

The tissue or biological sample can be fixed to a surface using fixatives. Preferred fixatives cause fixation of the cellular constituents through a precipitating action which is reversible, maintains a cellular morphology with the nucleic acids in the appropriate cellular location and does not interfere with nucleic acid hybridization. Fixatives include, for example, but are not limited to formaldehyde, alcohols, salt solutions, mercuric chloride, sodium chloride, sodium sulfate, potassium dichromate, potassium phosphate, ammonium bromide, calcium chloride, sodium acetate, lithium chloride, cesium acetate, calcium or

magnesium acetate, potassium nitrate, potassium dichromate, sodium chromate, potassium iodide, sodium iodate, sodium thiosulfate, picric acid, acetic acid, sodium hydroxide, acetones, chloroform, glycerin, and thymol.

After being fixed on a surface the samples are treated to remove proteins and other cellular material which may cause nonspecific background binding. Agents which remove protein include enzymes such as pronase or proteinase K, or mild acids, such 0.02-0.2N HCl. RNA can be removed with RNase.

The DNA on the surface must then be denatured so that the oligonucleotide probes can bind to give a signal. Denaturation can be accomplished by varying the pH, increasing temperature, or organic solvents such as formamide. The labeled probe is then hybridized with the denatured DNA under standard hybridization conditions.

The tissue or biological sample is any material which is composed of or contains cells or portions of cells. The cells may be living or dead and should contain substantially intact membranes sufficient to preserve the nucleic acid within the cell. The material may be deposited on the solid support using standard techniques such as sectioning of tissue or smearing or cytocentrifugation of single cell suspensions. Many types of solid supports may be used, including but not limited to, glass, nitrocellulose, scotch tape, nylon, or gene screen plus. A preferred support is a glass microscope slide.

A label as used herein is any molecule which may be detected. For instance, labels include but are not limited to ^{32}P, ^{14}C, ^{125}I, ^{3}H, ^{35}S biotin, avidin, fluorescent or enzymatic molecules. The nucleic acid may be labeled with biotin and then detected with avidin conjugated to a fluorescent, enzymatic, or colloidal gold conjugate. Biotin labeled nucleotides can be incorporated into the nucleic acid by nick translation, enzymatic, or chemical means.

In another aspect the invention is a probe for genomic subtractive hybridization. The probes of the mixture are isolated chemically modified oligonucleotide probes, wherein the chemically modified oligonucleotide probes are oligonucleotides associated with a target molecule selected from the group consisting of biotin, avidin, FITC, anti-FITC, antigen, and antibodies and wherein the oligonucleotide probes are a mixture of oligonucleotides complementary to repetitive nucleic acid sequences and yeast nucleic acid sequences.

An "isolated chemically modified oligonucleotide probe" is a chemically modified oligonucleotide probe as described above which has been substantially separated and purified

away from nucleic acid sequences in the cell of the organ in which the nucleic acid naturally occurs. The term isolated, thus, encompasses a chemically modified oligonucleotide probe wherein the oligonucleotide probe has a repetitive nucleic acid sequence and yeast DNA which is in a composition that is free of other oligonucleotides, including other unique oligonucleotides. The isolated chemically modified oligonucleotide probes, as used herein, encompasses a set of oligonucleotides having a mixture of repetitive nucleic acid sequences and yeast DNA, but does not encompass oligonucleotides having unique sequences.

Examples

Example 1: Preparation of Chromosomes.

Selection of YACs for MYC *and EWS Translocation Detection*

The *MYC* gene maps to chromosome subband 8q24.1 and is involved in three well characterized translocations in Burkitt lymphomas. The *EWS* gene maps to chromosome band 22q12 and is involved in at least eight different translocations in soft-tissue tumors. YAC clones centromeric and telomeric to these genes were selected based on published maps that were cross-referenced with physical mapping data from the Whitehead Institute Center for Genome Research and CEPH-Généthon websites (http://www.genome.wi.mit.edu and http://www.cephb.fr., respectively). YAC clones were selected based on appropriate map location and absence of features suggesting chimerism. Potential chimerism was determined using both sequence-lagged site (STS) and Alu-PCR data from the Whitehead Institute and CEPH. All YAC clones were obtained from Research Genetics (Huntsville, AL), and YAC DNAs were isolated as described previously. Chimerism was evaluated formally by fluorescence ISH (FISH) against normal male lymphocyte metaphase and interphase preparations and chimeric clones were excluded from the final contigs. Figure 1 depicts the YAC clones that comprise the centromeric and telomeric contigs flanking *MYC* and *EWS* on chromosomes 8 and 22.

Example 2: Preparation of Nucleic Acid Sample from Chromosomal DNA.

The nucleic acid sample (referred to herein as Tracer DNA) which is used to generate a library of probes for ISH discussed below was prepared by combining 2 μg of each pulsed field gel purified clone from a particular contig. These pooled DNAs were then sonicated to 0.1 to 8 kb and size-fractionated on 1.5% agarose gels. The 0.4- to 2-kb fractions were cut from the gels, purified using QlAquick gel extraction kits (QIAGEN, Santa Clarita, CA) blunt

ended and ligated to the T-1/T-2 adapter. The T-1/T-2 adapter was constructed by annealing polyacrylamide gel electrophoresis (PAGE)-purified oligos 5'CTGAGCGGAATTCGTGAGACC (T-1) SEQ. ID. NO: 1 and 5'GGTCTCACGAATTCCGCTCAGTT (T-2) SEQ. ID. NO: 2. Adapter-ligated fragments were then PCR amplified, in multiple 25 µl reactions, using the T-1 sequence as primer. Amplified fragments were purified using QIA-quick PCR purification kits (QIAGEN) and then eluted in 1 mmol/L Tris/Cl, pH 6.0. PCR reactions here and elsewhere, unless otherwise indicated, were done in 25 µl volumes using KlenTaq reaction buffer (Clontech, Palo Alto, CA). 0.2 mmol/L dNTPα, 1.2 µmol/L PAGE-purified primer, and 0.1 U/ml KlenTaq DNA polymerase (Clontech). PCR cycling conditions were 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 3 minutes for 25 to 30 cycles, followed by 72°C for 9 minutes.

Example 3: Preparation of Chemically Modified Oligonucleotide Probe.

The chemically modified oligonucleotide probes in the form of biotinylated repetitive oligonucleotide probes (referred to herein as Driver DNA) were prepared from three nonchimeric chromosome 21 YAC clones (746-b-10, 745-c-11, 615-c-9) that are known to be rich in repetitive sequences or from Cot-1 DNA which is enriched for repetitive DNA sequences. YAC DNA was isolated, as described, from a 500 ml culture consisting of all three clones. Nucleic acid thus isolated was precipitated twice in 6.5% polyethylene glycol/0.8 mol/L sodium chloride, washed in 70% ethanol, and dissolved in distilled water. Twenty micrograms of this DNA was sonicated and size fractionated as described above. Fragments ranging in size from 0.4 and 2 kb were gel purified, blunt ended, and ligated to a D-40/D-41 adapter constructed by annealing PAGE-purified oligos 5'AATTCTTGCGCCTTAAACCAAC (D-40) SEQ. ID. NO: 3 and 5'GTTGGTTTAAGGCGCAAG (D-41) SEQ. ID. NO: 4. PCR was performed using 5'(biotin)AATTCTTGCGCCTTAAACCAAC (D-40B) SEQ. ID. NO: 5 as primer. A second round of PCR was performed using the first-round product as the template. In the second round of PCR, the concentration of D-40B was increased to 6 mmol/L, the dNTP concentration was increased to 0.4 mmol/L, and the number of cycles was reduced to 20. Several hundred micrograms of biotinylated YAC repetitive oligonucleotide probes were generated in multiple 25 ml reactions. The biotinylated PCR products were purified using Qiaquick PCR purification kits (QIAGEN), precipitated in ethanol, and dissolved at 1.5

mg/ml in EE buffer (10 mmol/L 2hydroxyethyl]piperazine-N'-[3-propanesulfonic acid (NaEPPS), 1 mmol/L EDTA, pH 8.0).

Example 4: Subtractive Hybridization.

Genomic subtractive hybridization removes sequences from a tracer DNA population by hybridizing with a molar excess of driver DNA. The driver DNA is chemically modified, e.g., with a biotin, such that it may be selectively removed from solution along with driver-tracer hybrid molecules. Briefly, *MYC* and *EWS*-region YAC contigs (Figure 1) were repeatedly hybridized with a 40-fold excess of biotinylated YACs containing abundant repetitive, e.g., Alu and LINE element sequences. Consequently, repetitive sequences presenting the *EWS*- and *MYC*-region contigs were quantitatively removed. The detailed methods are set forth below.

Subtraction was performed by mixing 250 ng of tracer DNA with 10 μg of biotinylated driver DNA, 2 μg of T-1, 5 μg of yeast tRNA as carrier. This mixture was denatured at 99°C for 1 minute, lyophilized, redissolved in 5 μl of EE buffer/1 mol/L NaCl, and then incubated at 65°C for 24 to 48 hours. Biotinylated molecules (including tracer-driver hybrids) were removed using avidin-polystyrene beads as described. Remaining unbiotinylated tracer fragments were precipitated in ethanol before proceeding with the next round of subtraction. Each of three rounds of subtraction was performed exactly as described above. After the third round, remaining tracer fragments were amplified by PCR using the T-1 sequence as primer.

Example 5: In Situ Hybridization.

Probe Preparation

The subtracted contigs telomeric to the *EWS* and MYC loci (EWS.T and MYC.T, respectively, Figure 1) were labeled with biotin using the BioPrime random octamer priming kit (Gibco BRL/Life Technologies, Gaithersburg, MD). The subtracted contigs centromeric to the *EWS* and MYC loci (EWS.C and MYC.C, respectively, Figure 1) were labeled with fluorescein isothiocyanate (FITC), also by random octamer priming. The final nucleotide concentrations for FITC labeling were 0.2 mmol/L dCTP, 0.2 mmol/L dGTP, 0.2 mmol/L dATP, 0.1 mmol/L dTTP, and 0.1 mmol/L FITC-12-dUTP (NEN, Boston, MA). Residual primers and unincorporated nucleotides were removed by S-200HR spin-column chromatography (Pharmacia, Uppsala, Sweden). The purified products were precipitated in

ethanol and dissolved in a solution containing 50% formamide, 10% dextran sulfate, and 2X SSC (0.3 mol/l sodium chloride, 0.03 mol/L sodium citrate, pH 7.0).

Slide Preparation

Formalin-fixed, paraffin-embedded 4-μm tissue sections were applied to silanized slides, baked at 65°C for 16 hours, and stored at room temperature. Slides were processed for ISH using Oncor Tissue kits (Oncor, Gaithersburg, MD) according to the manufacturer's specifications with minor variations. Briefly, after deparaffination in xylene and dehydration in 100% ethanol, all tissue sections were incubated in 30% pretreatment solution for 15 minutes, followed by 15 to 40 minutes of protease treatment. Digestion times were optimized on a case-by-case basis. Alternately, paraffin sections were prepared for ISH using a combination of microwaving at 92°C followed by protein digestion using pepsin. Slides were denatured in a solution containing 70% formamide, 2X SSC, pH 7.0, at 75°C for 8 minutes. Slides were dehydrated in ice-cold 70%, 85%, and 95% ethanol and then air dried. Cytogenetic preparations were processed according to standard methods.

Hybridization and Detection

Aliquots of labeled DNA were diluted to a concentration of 100 to 200 ng/μl in hybridization solution (50% formamide, 10% dextran sulfate, 2X SSC), denatured at 75°C for 5 minutes, placed immediately onto denatured slides, and covered with a glass coverslip that was sealed with rubber cement. The slides were then placed in a humidified chamber at 37°C for 12 to 16 hours. Slides were washed in 0.5X SSC at 72°C for 5 minutes and then in PN buffer (0.1 mol/L sodium phosphate, pH 8.0, 0.1% Nonidet P-40) at room temperature. Biotinylated probes were detected with μg/ml Texas Red-strepavidin (Zymed Laboratories, South San Francisco, CA). FITC-labeled probes were visualized directly and, in some cases, amplified using rabbit anti-FITC and FITC anti-rabbit (Zymed). For colorimetric detection, sequential peroxidase reactions were performed using horseradish peroxidase (HRP)-conjugated goat anti-FITC (Zymed Laboratories, South San Francisco, CA) with the diaminobanzidine (DAB) substrate kit (Zymed), followed by HRP-conjugated strapavidin (Zymed) with the VIP substrate kit (Vector). Cells were counterstained with Gill's hematoxylin (Vector) and mounted in Permount (Sigma-Aldrich Corp., St. Louis, MO).

Dot Blotting

Total genomic DNA from *Saccharomyces cerevisae* strain AB1380 (a negative control) as well as subtracted and unsubtracted tracer DNA (EWS.C and EWS.T) were evaluated. Amounts of 300, 100, 30, 10, and 1 ng of each DNA were denatured and spotted onto a positively charged nylon membrane, which was then baked. The membrane was probed successively with radiolabeled human Cot-1 DNA (Gibco BRL/Life Technologies) and total yeast genomic DNA (strain AB1380).

Subtraction of Human and Yeast Repetitive Sequences

Dot blots of subtracted and unsubtracted tracer DNAs were evaluated for the presence of human repetitive sequences using human Cot-1 (repetitive-sequence-enriched) DNA as probe. This experiment demonstrated complete subtraction of human repetitive sequences from the tracer DNAs (Figure 2A). Reprobing with total yeast genomic DNA demonstrated substantial removal of yeast sequences (Figure 2B). All subtracted tracer DNA libraries were then evaluated as FISH probes, in the absence of Cot-1 competitor DNA or pre-annealing. FISH signals localized exclusively to the expected chromosome regions, and signal intensities were identical to those obtained using unsubtracted DNAs (the latter preannealed and hybridized in the presence of excess Cot-1 DNA to suppress nonspecific background staining). Potential PCR-related biases were evaluated by subjecting the subtracted probe pools to four successive rounds of PCR amplification. Probe aliquots were labeled after each round of amplification and were hybridized against two different slides. No dimunition in probe signal intensity was seen after four rounds of amplification. This experiment was validated by repeating the four rounds of PCR amplification using another aliquot of subtracted probe. Again, there was no fall-off FISH signal intensity after the fourth round of PCR. The Gibco BioPrime random octamer labeling approach permitted substantial amplification of the template DNA during the labeling process. Typical yields, starting with 200 ng of subtracted tracer, were 5-10 μg (25-60 fold amplification) of labeled DNA after a 5 hour, 50 μl random priming reaction. PCR incorporation, using the T-1 adapter primers with biotin- or digoxigenin-conjugated nucleotides, was an equally effective alternative to random priming.

Application 1: Evaluation of EWS-*Region Rearrangements*

The subtracted *EWS*-region ISH probe set was evaluated against 1) a primary cutaneous Ewing's sarcoma, 2) a clear-cell sarcoma, and 3) a tibial Ewing's sarcoma that lacked, cytogenetically, atypical (11:22). The cutaneous Ewing's sarcoma as reported previously, was an axillary lesion in a 19-year old woman. The clear-cell sarcoma was a right knee mass in a 30-year-old woman. The clear-cell sarcoma was a right knee mass in a 39-year-old woman. The histological differential diagnosis, for the knee mass, included both clear-cell sarcoma (melanoma of soft parts) and metastatic cutaneous melanoma. Greater than 65% of clear-cell carcomas contain a (12:22)(q13.q12), resulting in fusion of *EWS* and *ATF-1*[27] whereas this translocation has not been reported in cutaneous malignant melanoma. FISH analysis of 4 μm paraffin sections revealed splitting of one EWS.C/EWS.T probe pair, consistent with *EWS*-region rearrangement, in both the cutaneous Ewing's sarcoma and putative clear-cell sarcoma. The tibial Ewing's sarcoma, diagnosed in a 17-year-old boy, had a cytogenetically aberrant chromosome 22 homolog, whereas chromosomes 2, 7, 11, 17, and 21 (chromosomes containing ETS family genes involved in known Ewing's sarcoma *EWS* fusions) were unremarkable by banding analysis. EWS.C/EWS.T FISH evaluation revealed a reciprocal translocation (1:22)(q42:q12), of the *EWS* region, and reverse transcriptase PCR was negative for *EWS-FLI1* or *EWS-ERG* fusion transcripts. These data support a unique *EWS* translocation, potentially involving a novel ETS family locus on chromosome band 1q42.

Application 2: Evaluation of MYC-*Region Rearrangement*

The subtracted *MYC*-region ISH probe was evaluated against a malignant pleural effusion in which cytological evaluation, but not immunophenotype, was classical for Burkill's lymphoma. The pleural fluid was from a 10-year-old boy with a 2-month history of anorexia and abdominal pain and with radiolocal evidence of mesenteric/mediastinal adenopathy and bilateral pleural effusions. Cytological evaluation revealed a homogeneous population of small lymphoid cells with bluish vacuolated cytoplasm and round nongrooved nuclei, whereas flow immunophenotyping was notable for the absence of surface immunoglobulin. Cytogenetic banding studies were inconclusive because the chromosome morphology was poor. FISH analysis, using the subtracted MYC.C/MYC.T probe set demonstrated *MYC*-region rearrangement in metaphase and interphase cells. *MYC*-region rearrangement was also demonstrated convincingly by colorimetric detection.

In summary, we report a new method for constructing and synthesizing DNA probes from multimegabase YAC contigs. This method enables creation of adapter-ligated DNA libraries that are free of certain sequences such as repetitive sequences. We demonstrate that subtracted unique-sequence probes are detected readily using standard fluorescence and calorimetric reagents. The probes are labeled conveniently and are hybridized without competitor DNA or pre-annealing, thus simplifying the *In situ* hybridization protocol.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by examples provided, since the examples are intended as a single illustration of one aspect of the invention and other functionally equivalent embodiments are within the scope of the invention. Various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims. The advantages and objects of the invention are not necessarily encompassed by each embodiment of the invention.

All references, patents and patent publications that are recited in this application are incorporated in their entirety by reference.

References

1. Rabbitts TH: Chromosomal translocations in human cancer. Nature 1994, 372:143-149

2. Mitelman F, Mertens F, Johansson B: A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nat Genet 1997, 15 Spec No:417-474

3. Look AT: Oncogenic transcription factors in the human acute leukemias. Science 1997, 278:1059-1064

4. Delattre O, Zucman J, Melot T, Garau XS, Zucker JM, Lenoir GM, Ambros PF, Sheer D, Turc-Carel C, Triche TJ, Aurias A, Thomas G: The Ewing family of tumors--a subgroup of small-round-cell tumors defined by specific chimeric transcripts. N Engl J Med 1994, 331:294-299

5. Sandberg AA, Turc-Carel C, Gemmill RM: Chromosomes in solid tumors and beyond. Cancer Res 1988, 48:1049-1059

6. Sreekantaiah C, Ladanyi M, Rodriguez E, Chaganti RS: Chromosomal aberrations in soft tissue tumors. Relevance to diagnosis, classification, and molecular mechanisms. Am J Pathol 1994, 144:1121-1134

7. Speel EJ, Ramaekers FC, Hopman AH: Cytochemical detection systems for in situ hybridization, and the combination with immunocytochemistry, 'who is still afraid of red, green and blue?'. Histochem J 1995, 27:833-858

8. Hopman AH, Claessen S, Speel EJ: Multi-colour brightfield in situ hybridisation on tissue sections. Histochem Cell Biol 1997, 108:291-298

9. Hopman AH, Poddighe PJ, Smeets AW, Moesker O, Beck JL, Vooijs GP, Ramaekers FC: Detection of numerical chromosome aberrations in bladder cancer by in situ hybridization. Am J Pathol 1989, 135:1105-1117

10. Cremer T, Lichter P, Borden J, Ward DC, Manuelidis L: Detection of chromosome aberrations in metaphase and interphase tumor cells by in situ hybridization using chromosome-specific library probes. Hum Genet 1988, 80:235-246

11. Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, Gray J: Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A 1988, 85:9138-9142

12. Tkachuk DC, Westbrook CA, Andreeff M, Donlon TA, Cleary ML, Suryanarayan K, Homge M, Redner A, Gray J, Pinkel D: Detection of bcr-abl fusion in chronic myelogeneous leukemia by in situ hybridization. Science 1990, 250:559-562

13. Xiao S, Renshaw AA, Cibas ES, Hudson TJ, Fletcher JA: Novel fluorescence in situ hybridization approaches in solid tumors: Characterization of frozen specimens, touch preparations, and cytologic preparations. Am J Pathol 1995, 147:896-904

14. Poddighe PJ, Moesker O, Smeets D, Awwad BH, Ramaekers FC, Hopman AH: Interphase cytogenetics of hematological cancer: comparison of classical karyotyping and in situ hybridization using a panel of eleven chromosome specific DNA probes. Cancer Res 1991, 51:1959-1967

15. Gray JW, Pinkel D: Molecular cytogenetics in human cancer diagnosis. Cancer 1992, 69:1536-1542

16. Wolman SR: Fluorescence in situ hybridization: a new tool for the pathologist. Hum Pathol 1994, 25:586-590

17. Tiainen M, Hopman A, Moesker O, Ramaekers F, Wessman M, Laasonen A, Pyrhonen S, Tammilehto L, Mattson K, Knuutila S: Interphase cytogenetics on paraffin sections of

malignant pleural mesothelioma. A comparison to conventional karyotyping and flow cytometric studies. Cancer Genet Cytogenet 1992, 62:171-179

18. Manuelidis L, Langer-Safer PR, Ward DC: High-resolution mapping of satellite DNA using biotin-labeled DNA probes. J Cell Biol 1982, 95:619-625

19. Pinkel D, Straume T, Gray JW: Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Natl Acad Sci U S A 1986, 83:2934-2938

20. Lichter P, Tang CJ, Call K, Hermanson G, Evans GA, Housman D, Ward DC: High-resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones. Science 1990, 247:64-69

21. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D: Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 1992, 258:818-821

22. Speicher MR, Gwyn Ballard S, Ward DC: Karyotyping human chromosomes by combinatorial multi-fluor FISH. Nat Genet 1996, 12:368-375

23. Schrock E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y, Ried T: Multicolor spectral karyotyping of human chromosomes. Science 1996, 273:494-497

24. Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S, Leder P: Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A 1982, 79:7837-7841

25. Davis M, Malcolm S, Rabbitts TH: Chromosome translocation can occur on either side of the c-myc oncogene in Burkitt lymphoma cells. Nature 1984, 308:286-288

26. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, Aurias A, Thomas G: Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature 1992, 359:162-165

27. Zucman J, Delattre O, Desmaze C, Epstein AL, Stenman G, Speleman F, Fletcher CD, Aurias A, Thomas G: EWS and ATF-1 gene fusion induced by t(12;22) translocation in malignant melanoma of soft parts. Nat Genet 1993, 4:341-345

28. Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT: A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. Nat Genet 1994, 6:146-151

29. Ladanyi M, Gerald W: Fusion of the EWS and WT1 genes in the desmoplastic small round cell tumor. Cancer Res 1994, 54:2837-2840

30. Jeon IS, Davis JN, Braun BS, Sublett JE, Roussel MF, Denny CT, Shapiro DN: A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1. Oncogene 1995, 10:1229-1234

31. Gill S, McManus AP, Crew AJ, Benjamin H, Sheer D, Gusterson BA, Pinkerton CR, Patel K, Cooper CS, Shipley JM: Fusion of the EWS gene to a DNA segment from 9q22-31 in a human myxoid chondrosarcoma. Genes Chromosom Cancer 1995, 12:307-310

32. Clark J, Benjamin H, Gill S, Sidhar S, Goodwin G, Crew J, Gusterson BA, Shipley J, Cooper CS: Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma. Oncogene 1996, 12:229-235

33. Panagopoulos I, Hoglund M, Mertens F, Mandahl N, Mitelman F, Aman P: Fusion of the EWS and CHOP genes in myxoid liposarcoma. Oncogene 1996, 12:489-494

34. Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H, Delattre O: A new member of the ETS family fused to EWS in Ewing tumors. Oncogene 1997, 14:1159-1164

35. Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH, Hu S, Colbert AME, Rosenberg C, Reeve-Daly MP, Rozen S, Hui L, Wu X, Vestergaard C, Wilson KM, Bae JS, Maitra S, Gaiatsas S, Evans CA, DeAngelis MM, Ingalls KA, Nahf RW, Horton LT, Anderson MO, Collymore AJ, Ye W, Kouyoumijian V, Zemsteva IS, Tarn J, Devine R, Courtney DF, Renaud MT, Nguyen H, O'Connor TJ, Fizames C, Faure S, Gyapay G, Dib C, Morissette J, Orlin JB, Birren BW, Goodman N, Weissenbach J, Hawkins TL, Foote S, Page DCet al.An STS-based map of the human genome. Science 1995, 270:1945-1954

36. Collins JE, Cole CG, Smink LJ, Garrett CL, Leversha MA, Soderlund CA, Maslen GL, Everett LA, Rice KM, Coffey AJ, Gregory SG, Gwilliam R, Dunham A, Davies AF, Hassock S, Todd CM, Lehrach H, Hulsebos TJM, Weissenbach J, Morrow B, Kucherlapati RS, Wadey R, Scambler PJ, Kim UJ, Simon MI, Peyrard M, Xie YG, Carter NP, Durbin R,

Dumanski JP, Bentley DR, Dunham I: A high-density YAC contig map of human chromosome 22. Nature 1995, 377:367-379

37. Kim UJ, Shizuya H, Kang HL, Choi SS, Garrett CL, Smink LJ, Birren BW, Korenberg JR, Dunham I, Simon MI: A bacterial artificial chromosome-based framework contig map of human chromosome 22q. Proc Natl Acad Sci U S A 1996, 93:6297-6301

38. Huang SF, Xiao S, Renshaw AA, Loughlin KR, Hudson TJ, Fletcher JA: Fluorescence in situ hybridization evaluation of chromosome deletion patterns in prostate cancer. Am J Pathol 1996, 149:1565-1573

39. Milner JJ, Cecchini E, Dominy PJ: A kinetic model for subtractive hybridization. Nucleic Acids Res 1995, 23:176-187

40. Lovett M: Direct selection of cDNAs using genomic contigs. In: Current Protocols in Human Genetics. Dracopoli NC, Haines JL, Korf BR, et al, John Wiley and Sons, Inc. 1994; 6.3.1-6.3.15.

41. Xu H, Wei H, Tassone F, Graw S, Gardiner K, Weissman SM: A search for genes from the dark band regions of human chromosome 21. Genomics 1995, 27:1-8

42. Silverman GA: Purification of YAC-containing total yeast DNA. In: YAC Protocols. Markie D, ed. Totawa, NJ: Humana Press Inc. 1995; 65-68.

43. Straus D, Ausubel FM: Genomic subtraction for cloning DNA corresponding to deletion mutations. Proc Natl Acad Sci U S A 1990, 87:1889-1893

44. Lanciego JL, Goede PH, Witter MP, Wouterlood FG: Use of peroxidase substrate Vector VIP for multiple staining in light microscopy. J Neurosci Methods 1997, 74:1-7

45. Mackey J, Darfler M, Nisson P, Rashtchian A: Use of random primer extension for concurrent amplification and nonradioactive labeling of nucleic acids. Anal Biochem 1993, 212:428-435

46. Hasegawa SL, Davison JM, Rutten A, Fletcher JA, Fletcher CD: Primary cutaneous Ewing's sarcoma: immunophenotypic and molecular cytogenetic evaluation of five cases. Am J Surg Pathol 1998, 22:310-318

47. Nedivi E, Hevroni D, Naot D, Israeli D, Citri Y: Numerous candidate plasticity-related genes revealed by differential cDNA cloning. Nature 1993, 363:718-722

48. Mor O, Messinger Y, Rotman G, Bar-Am I, Ravia Y, Eddy RL, Shows TB, Park JG, Gazdar AF, Shiloh Y: Novel DNA sequences at chromosome 10q26 are amplified in human

gastric carcinoma cell lines: molecular cloning by competitive DNA reassociation. Nucleic Acids Res 1991, 19:117-123

49. Nussbaum RL, Lesko JG, Lewis RA, Ledbetter SA, Ledbetter DH: Isolation of anonymous DNA sequences from within a submicroscopic X chromosomal deletion in a patient with choroideremia, deafness, and mental retardation. Proc Natl Acad Sci U S A 1987, 84:6521-6525

50. Kunkel LM, Monaco AP, Middlesworth W, Ochs HD, Latt SA: Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. Proc Natl Acad Sci U S A 1985, 82:4778-4782

51. Lamar EE, Palmer E: Y-encoded, species-specific DNA in mice: evidence that the Y chromosome exists in two polymorphic forms in inbred strains. Cell 1984, 37:171-177

52. Rosenberg M, Przybylska M, Straus D: "RFLP subtraction": a method for making libraries of polymorphic markers. Proc Natl Acad Sci U S A 1994, 91:6113-6117

53. Reijo R, Lee TY, Salo P, Alagappan R, Brown LG, Rosenberg M, Rozen S, Jaffe T, Straus D, Hovatta O, et al: Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nat Genet 1995, 10:383-393

54. Guan XY, Meltzer PS, Trent JM: Rapid generation of whole chromosome painting probes (WCPs) by chromosome microdissection. Genomics 1994, 22:101-107

55. Craig JM, Kraus J, Cremer T: Removal of repetitive sequences from FISH probes using PCR-assisted affinity chromatography. Hum Genet 1997, 100:472-476

56. Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC: Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet 1988, 80:224-234

57. Kohno K, Oshiro T, Kishine H, Wada M, Takeda H, Ihara N, Imamoto F, Kano Y, Schlessinger D: Construction and characterization of a rad51rad52 double mutant as a host for YAC libraries. Gene 1997, 188:175-181

58. Le Y, Dobson MJ: Stabilization of yeast artificial chromosome clones in a rad54-3 recombination-deficient host strain. Nucleic Acids Res 1997, 25:1248-1253

59. Lengauer C, Green ED, Cremer T: Fluorescence in situ hybridization of YAC clones after Alu-PCR amplification. Genomics 1992, 13:826-828

60. Lengauer C, Riethman HC, Speicher MR, Taniwaki M, Konecki D, Green ED, Becher R, Olson MV, Cremer T: Metaphase and interphase cytogenetics with *Alu*-PCR-amplified yeast artificial chromosome clones containing the *BCR* gene and the protooncogenes *c-raf-1*, *c-fms*, and *c-erbB-2*. Cancer Res 1992, 52:2590-2596

61. Bohlander SK, Espinosa R, 3d, Fernald AA, Rowley JD, Le Beau MM, Diaz MO, Espinosa R: Sequence-independent amplification and labeling of yeast artificial chromosomes for fluorescence in situ hybridization. Cytogenet Cell Genet 1994, 65:108-110

62. Guan XY, Trent JM, Meltzer PS: Generation of band-specific painting probes from a single microdissected chromosome. Hum Mol Genet 1993, 2:1117-1121

63. Joos S, Haluska FG, Falk MH, Henglein B, Hameister H, Croce CM, Bornkamm GW: Mapping chromosomal breakpoints of Burkitt's t(8;14) translocations far upstream of c-myc. Cancer Res 1992, 52:6547-6552

64. Veronese ML, Ohta M, Finan J, Nowell PC, Croce CM: Detection of myc translocations in lymphoma cells by fluorescence in situ hybridization with yeast artificial chromosomes. Blood 1995, 85:2132-2138

65. Pelicci PG, Knowles DM, 2d, Magrath I, Dalla-Favera R, Knowles DM: Chromosomal breakpoints and structural alterations of the c-myc locus differ in endemic and sporadic forms of Burkitt lymphoma. Proc Natl Acad Sci U S A 1986, 83:2984-2988

66. Zeidler R, Joos S, Delecluse HJ, Klobeck G, Vuillaume M, Lenoir GM, Bornkamm GW, Lipp M: Breakpoints of Burkitt's lymphoma t(8;22) translocations map within a distance of 300 kb downstream of MYC. Genes Chromosom Cancer 1994, 9:282-287

We claim:

CLAIMS

1. A method for genomic subtractive hybridization, comprising:

hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in a nucleic acid sample, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule, and

selectively removing the chemically modified oligonucleotide probe and complementary nucleic acid by selectively contacting the target molecule with a binding partner to produce binding partner/target conjugates and separating the binding partner and binding partner/target conjugates from the nucleic acid sample.

2. The method of claim 1, wherein the target molecule is selected from the group consisting of avidin, biotin, FITC, anti-FITC, antigen, and antibodies.

3. The method of claim 1, wherein the oligonucleotide is a repetitive nucleic acid.

4. The method of claim 3, wherein the repetitive nucleic acid is isolated from a DNA source selected from the group consisting of YAC, BAC and PAC DNA.

5. The method of claim 3, wherein the chemically modified oligonucleotide probe is prepared by amplifying the repetitive nucleic acid using PCR and a primer attached to the target molecule.

6. The method of claim 1, wherein the nucleic acid sample is a genomic nucleic acid sample.

7. The method of claim 1, wherein the genomic nucleic acid sample is obtained from a YAC clone.

8. The method of claim 7, wherein the YAC clone is purified by pulsed field gel electrophoresis.

9. The method of claim 1, wherein the binding partner is immobilized on a support.

10. The method of claim 9, wherein the support is a bead and wherein the binding partner/target conjugates are separated from the nucleic acid sample by chromatography.

11. The method of claim 1, wherein the genomic nucleic acid sample is obtained from a DNA source selected from the group consisting of YAC, BAC and PAC DNA.

12. A library of nucleic acid probes for use in *in situ* hybridization methods, comprising:

a heterogenous mixture of labeled nucleic acid probes that are substantially complementary to unique nucleic acid fragments and are substantially free of repetitive nucleic acid sequences, and which are produced by the process of:

(a) obtaining genomic nucleic acid fragments;

(b) amplifying the genomic nucleic acid fragments;

(c) hybridizing a chemically modified oligonucleotide probe to a complementary nucleic acid sequence in the genomic nucleic acid fragments, wherein the chemically modified oligonucleotide probe is an oligonucleotide associated with a target molecule;

(d) selectively removing the chemically modified oligonucleotide probe and complementary nucleic acid by selectively contacting the target molecule with a binding partner and separating the binding partner and binding partner/target conjugates from the genomic nucleic acid fragments; and

(e) labeling the genomic nucleic acid fragments with a label.

13. The library of claim 12, wherein the genomic nucleic acid fragments are labeled with a fluorescent molecule.

14. The library of claim 12, wherein the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of a genomic DNA population.

15. The library of claim 12, wherein the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of a chromosome of a genomic DNA population.

16. The library of claim 12, wherein the heterogenous mixture of labeled nucleic acid probes are complementary to substantially all of a subregion of a chromosome of a genomic DNA population.

17. A method for performing in situ hybridization, comprising:

hybridizing the library of labeled probes of claim 12 to a biological sample fixed on a surface,

removing un-hybridized probe, and

detecting a signal from the hybridized probe to identify nucleic acid sequences present in the biological sample.

Figure 1

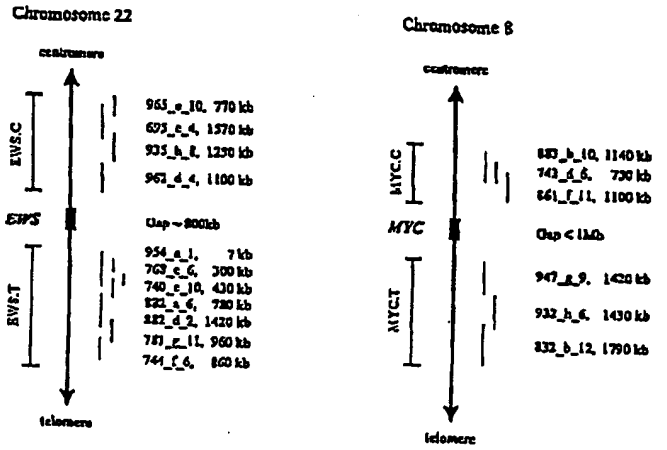
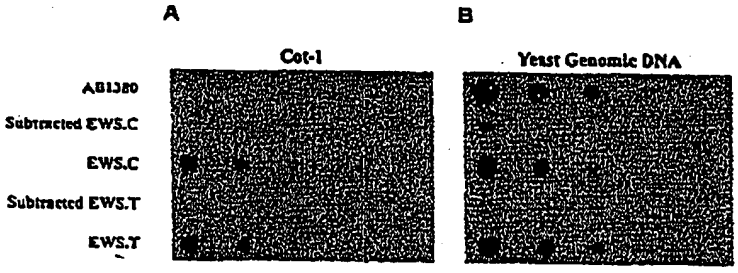

Chromosome 22
centromere
EWS.C
EWS
EWS.T
telomere
963_e_10, 770 kb
693_c_4, 1570 kb
935_h_8, 1250 kb
967_d_4, 1100 kb
Gap ~ 800kb
954_a_1, 7 kb
769_c_6, 300 kb
740_c_10, 430 kb
822_a_6, 720 kb
882_d_2, 1420 kb
781_g_11, 960 kb
744_f_6, 860 kb
Chromosome 8
centromere
MYC.C
MYC
MYC.T
telomere
883_b_10, 1140 kb
743_f_6, 730 kb
661_f_11, 1100 kb
Gap < 1Mb
947_g_9, 1420 kb
932_h_6, 1430 kb
832_b_12, 1790 kb

Figure 2


A
B
Cot-1
Yeast Genomic DNA
AB1380
Subtracted EWS.C
EWS.C
Subtracted EWS.T
EWS.T

SEQUENCE LISTING

<110> THE BRIGHAM AND WOMEN'S HOSPITAL, INC.
FLETCHER, Jonathan
XIAO, Sheng

<120> A METHOD FOR GENOMIC SUBTRACTIVE HYBRIDIZATION

<130> Z0093/7001/HCL

<150> U.S.S.N. 60/106,701
<151> 1998-11-02

<160> 5

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> HOMO SAPIENS

<400> 1
ctgagcggaa ttcgtgagac c 21

<210> 2
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> HOMO SAPIENS

<400> 2
ggtctcacga attccgctca gtt 23

<210> 3
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> HOMO SAPIENS

<400> 3
aattcttgcg ccttaaacca ac 22

<210> 4

<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> HOMO SAPIENS

<400> 4
gttggtttaa ggcgcaag 18

<210> 5
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> HOMO SAPIENS

<400> 5
aattcttgcg ccttaaacca ac 22